Faculty Relevance Criteria: Internalized User Needs

Lynn Westbrook

Abstract

Reference librarians, online searchers, system designers, and other information professionals work to incorporate user-based relevance criteria into information services and systems. Genuinely utilizing the relevance criteria that patrons employ requires, as a first step, the development of an in-depth understanding of those criteria. This study progresses toward that understanding by providing new data on the criteria used by members of a rarely studied interdisciplinary population and by developing a user-centered methodology. Each of five Women's Studies faculty members was interviewed concerning her immediate reactions to different resources provided in direct response to real, on-going information needs. The criteria identified by this approach went beyond topicality, currency, and other basic aspects of relevance criteria to include elements such as interdisciplinarity and theoretical perspective.

Introduction

The concept of relevance remains central to library and information studies work, the gold standard by which services and systems are judged. Those that put relevant information into patrons' hands succeed; all others fail to some extent. The practical aspects of relevance work may be loosely divided into two approaches, system-centered and user-centered. The system-centered approach seeks to develop the competence of an external agency (such as a database or search engine) that can recognize relevant information by its similarities to the semantic representation of an information need. A series of OPAC searches, for example, should retrieve rel-

Lynn Westbrook, School of Library and Information Studies, Texas Woman's University, P.O. Box 425438, Denton, TX 76204–3438
LIBRARY TRENDS, Vol. 50, No. 2, Fall 2001, pp. 197–206

evant items on the basis of semantic clues in the search statement (e.g., words used, context of word use, and word relationships) by placing them in relationship to the document representation (i.e., the cataloging record). The system-centered work builds on an assumption that the information embodied in a document or its surrogate can be objectively, if somewhat imprecisely, matched to the actual information need as articulated in a structured query (Mizzaro, 1997, p. 812). This approach leads to developments in document representation and information retrieval (e.g., Saracevic, 1969; Bookstein, 1979; Janes, 1991).

The user-centered approach to relevance work holds that such a complex, highly personalized array of factors comprise relevance for a single individual in a certain situation at any given point in time—that no outsider can accurately judge relevance. Reference librarians often observe this phenomenon in action when users reject a number of "relevant" items in favor of what appears to be an irrelevant item. This user-centered approach leads to a taxonomy of user-generated criteria and an enhanced understanding of the elements involved in making relevance judgments (e.g., Cool, Belkin, & Kantor, 1993; Park, 1993; Barry, 1994). Reference librarians who understand the full range of possible relevance criteria can effectively serve their patrons. Both approaches strengthen the effort to link individuals with the information they need.

Research Question

Contributing to the user-centered approach, this study seeks to identify and describe the criteria that members of one population (faculty teaching Women's Studies courses) apply to information provided, at their request, to assist them in meeting one type of information need (a single curriculum-development project). What range of criteria is possible? This study characterizes and classes the factors pertinent to this sample of this user group in this setting.

Since so little research has been conducted on the criteria used by interdisciplinary scholars, and since nothing at all has been done on Women's Studies faculty, no hypothesis is yet ready for testing; therefore this is an exploratory study. While no generalizations are claimed on the basis of these data, a range of criteria have been identified and delineated. Finally, some tentative explanations are proposed for the consideration of reference librarians, Women's Studies faculty, and Library and Information Studies scholars.

Data-Gathering Methodology

In this study, the relevance definition centered on "utility" (see Regazzi, 1988). As in natural information-seeking situations, participants made their own determinations regarding the effort worth expending on each item. In this endeavor to understand what makes information appear valuable, useful, worth some effort to obtain, or simply "good to know about," it was

critical that the methodological approach focus on actual needs of the participants. A qualitative approach was employed, using an unstructured interview as the primary data-gathering technique and content analysis as the primary data-analysis technique.

For the purposes of this small-scale exploratory study the population was defined as faculty (tenure-track and lecturers, but not teaching assistants) who were actively interested or engaged in a curriculum development project for a course offered or cross-listed by the Women's Studies program of a large midwestern university. As Women's Studies is a highly interdisciplinary field, the two participant-recruitment methods were designed to maximize representation of this diversity. First, the researcher personally invited three faculty members whose teaching areas represented varied information needs: an international social policy perspective, a medical science perspective, and a fine arts perspective. Second, an electronic notice soliciting participation was sent to the Women's Studies chair, who forwarded it with her recommendation to all twenty-three members of the Women's Studies faculty. Two more participants responded electronically and, after brief telephone conversations, elected to join the study. Their information needs represented racial, historical, and lesbian studies perspectives.

Demographic variables were not sought from participants on the grounds that the sample was far too small to provide meaningful information. It is worth noting, however, that Joanna Lewis is African American and probably the youngest participant; Kate Jacobs and Frankie Taylor are a generation older than the others; Jo Lawler is probably not American, given her accent; and Margaret Goodman had a child in daycare at the time of this study, a fact which may have heightened her interest in her area of study. (All participant names are pseudonyms.)

After each participant signed her consent form, an initial interview was scheduled in order to develop as full an understanding as possible of her information needs. The interviews were, with the permission of the participants, taped. Complete transcriptions of these tapes provided some data.

Participants were asked to explain the course and what they needed for it as if they were talking to a research assistant (RA) who would then go out and find the information for them. (While Lewis had not yet had an RA, the others were accustomed to that relationship. Without exception each participant explicitly stated that the researcher was viewed as an experienced librarian, and not as an RA, in that the need was explained in greater depth than it would be for an RA.) As they explained what was needed, probing and clarifying questions were used to elicit additional detail. When nothing more was forthcoming, the researcher asked follow-up questions on the basic points commonly covered in an extensive reference interview; e.g., geographical limitations, language, and preferred information format. In addition, the participants were questioned about their preferences regarding the format of the research results; e.g., printouts of citations alone or citations with

search terms. Finally, they were asked to specify their preference for the format of the meeting at which the search results would be presented to them. They were asked to work in the way they found most comfortable and typical, so long as it allowed face-to-face interaction at some point. Four of the five wanted to review searches with their RAs in person and the fifth was willing to do so. For all five participants computer printouts were quite acceptable. No firm limitations as to the number of cites were set.

In this interview the details of the verbal contract between participant and researcher were firmly established. The faculty were to state a real information need and to offer informal verbal feedback on the results of the searches done for them. The researcher was to provide the best possible search results and was allowed to include items about which there was a sincere question as to relevance. No set number of interviews was requested; the determining factor would be the extent of the participants' interest in continuing to review the information gathered in response to their stated needs.

Based on the understanding developed in the initial interview, the information search was conducted as soon as possible. A variety of resources (print, CD, and online) were used in an effort to find whatever would best meet the need. (Web sites were not of interest to any of the participants at the time of this study; given the rapid growth of sites in Women's Studies, however, that preference may well have changed.) Common tools included: the university online catalog, several Wilson indexes, *Dissertation Abstracts International* on CD-ROM, and the *Social Science Citation Index* on CD-ROM. Resulting citations were printed out with abstracts whenever possible. Pages from reference books were photocopied. The final pool of results was then grouped into a single stack with each separate citation numbered. (Dr. Jacobs eventually requested that certain information be created for her, such as lists of United Nations agencies and the most useful OPAC subject headings. Each of these lists was counted as a single item.) The results were given to the participant at each interview with no copy kept by the researcher. Since the purpose was to understand the breadth and range of their reactions, there was no need to record reactions to specific items. Over 1,000 items were presented to the five participants during the course of the study, as indicated in Table 1.

Table 1: Number of Cites per Interview

	1	2	3	4	Total
Lewis	114	89	80	—	283
Jacobs	52	9	16	5	82
Lawler	55	67	58	2	182
Goodman	150	125	—	—	275
Taylor	65	34	63	91	253
Totals	436	324	217	98	1075

After the searches were completed, reaction interviews were held by appointment. None of them had a preset length; the time needed for reviewing and discussing the results was always available.

The structure for each interview was the same. First, the search result sets were presented, with a sentence or two to describe the range of sources consulted. The participant would usually choose to begin the interview right away. (Occasionally a general, silent skimming of the items was done first; in one case the entire list was reviewed carefully before beginning.) Second, the participant would, with pen in hand, go through the list item by item. In addition to audio-taping, the researcher took written notes of reactions to the information and asked questions to clarify responses as needed while the participant made notes according to her personal system. Third, as contradictions seemed to arise between one reaction and another, the researcher stopped the review of results, when possible, just long enough to ask for a clarification. Fourth, as new, expanded, or more specific aspects of the need were expressed, the researcher took note of them. Finally, the researcher verbally summarized what was still needed and offered new hints regarding what was not wanted. The summary was posed as a question so that the participant was encouraged to augment, explain, or correct any part of it.

Throughout these interviews a great effort was made to encourage participants to feel comfortable about and to fully express their negative reactions. Since there might be a natural tendency to respond positively to the sheer effort made on their behalf, an emphasis was placed on the great value of negative reactions. Thus, criteria relating to both what was to be sought (e.g., authoritative works) and what was to be avoided (e.g., case studies) were identified.

Finally, member-checking interviews were held after the initial data analysis was complete. Each participant was contacted as soon as possible after the last reaction interview to set up a final, brief meeting. At that time they were given a copy of the summary notes on the general characteristics of their information criteria. They were given an opportunity to provide feedback and further information on their criteria as understood by the researcher.

Data Analysis

For each participant, the coding of the initial interview was completed before the first reaction interview took place. Using the constant comparison method of coding (Glaser, 1965), the transcript of the interview tape was reviewed phrase by phrase. Starting without preconceived categories, the researcher identified and eventually defined the categories into which selected phrases fell. Working through successive interviews and concentrating only on those phrases that described the information need and criteria, the researcher grouped similar items. When new categories were

formed, everything that had already been coded was reviewed again, to insure that nothing had been left out. Throughout this reiterative process, categories were formed, reformed, split, and combined until a final set of codes became consistently useful. At this point the final definitions were written and final checks made to confirm complete, accurate, and consistent coding.

Trustworthiness and Validity Checks

The primary validity check of the data gathering was the member-checking interview, designed to identify missed or misinterpreted data. Each participant agreed that the criteria identified were in use, that none was missing, and that none had been misidentified. They noted that the exact same criteria would not apply to all of their curricular work, but that several elements were stable. Each participant mentioned keeping the criteria list handy as a means of helping future RAs to understand their information needs.

The primary validity check in the data analysis involved recoding. Random sections of each interview were moved to a clean file and recoded using the established categories. Coding decisions matched with a 90% accuracy rate, and no new categories were formed. Only after this level of accuracy has been reached did the final analysis take place.

Findings

The coded material broke down into five groups. One group of codes included established relevance components that are commonly covered in indexing schemes; another included established components not commonly covered. A third included relatively unexplored relevance criteria. The final two groups concerned items related to the decision-making process and items related to the research study itself. While all of this material is of interest, the first three groups are of particular importance.

Seven relevance components appeared quickly, as expected. These common elements are already covered in most indexing schemes: topic, subtopic, currency, geographical parameters, temporal parameters, language, and length. Long acknowledged as patterns in classification and indexing schemes, these elements also appear in the literature of online and reference interview search strategies.

The second group of codes concerned relevance criteria that are commonly acknowledged in the reference interview but problematic in indexing schemes. This group includes the amount of material needed, its availability, the citation format, the primary or secondary nature of the material, and its general quality. These items differ in nature; the first one (amount of material needed) applies to the search as a whole, while all of the others apply to a single item. Availability is sometimes indicated on various systems, but nothing is accurate at the shelf level. Citation format is increas-

ingly flexible but still far from standardized. Primary and secondary materials are sometimes indicated by terms such as "diaries" or "essays," but they too are not represented fully. Of course, some people determine information quality by examining source (e.g., the title of a refereed journal), publisher (e.g., a university press), or author (e.g., a known expert). Quality as a characteristic is not, of course, objectively determined and is not, therefore, likely to be incorporated into an indexing scheme.

In addition to these expected elements of relevance, another group of criteria emerged. These twelve criteria are not universally acknowledged in either system design (e.g., classification schemes and subject terminology) or system interface (e.g., reference interviews and search strategies). They are, however, in regular use by the participants in this case study and divide into three groups: those that are internal to the participant, those that develop from the teaching focus, and those that characterize information.

The purely internal tie directly to the life experience and personal perspective of the participant and are, therefore, difficult to recognize and predict. The four criteria grouped in this area are curiosity, personal interest, redundancy, and personal knowledge.

On rare occasions, an item piqued the curiosity of individuals without particular reference to the immediate issue at hand. Dr. Lawler, for example, noted that an author had made the TV talk-show circuit with a book; she wanted to see it just to find out "why it's so popular," even though she would not otherwise have been interested enough to pursue it (Interview 2). All of her other relevance criteria would have called for the rejection of that work but curiosity kept it on the list.

Similarly, personal interest in an item might arise out of its use in another setting. The interest could come from work on another course, work on a research project, or something completely private. Dr. Lewis, for example, said, "I need to see that for my own needs" when she was thinking of writing a similar piece and submitting it to that same journal. Her goal was "to see if I'm wasting my time" (Interview 2).

While the first two of these four criteria kept some items in the pool, redundancy excluded items from consideration. Dr. Taylor's comment regarding redundancy was typical: "This doesn't tell me anything I don't already know" (Interview 3). Weeding out duplicates failed to eliminate redundant items, especially for those faculty who had been teaching in an area for some time.

Personal knowledge, however, was a two-edged sword. Personal familiarity with an author, conference, or journal could keep an item in or out of consideration, depending on the nature of the experience. Dr. Goodman, for example, had little interest in British publications on a particular topic because she'd found the scholarship so inadequate in the past. Lacking personal knowledge did not, however, relegate an item to obscurity.

These faculty repeatedly chose to seek out articles from unfamiliar journals so they could judge the journal quality for themselves. They did not assume that an unfamiliar journal was either outside their sphere of interest or of poor quality.

As might be expected when working with faculty who are involved in course development, some criteria tied directly to teaching, readability, and engendering a response in students. Each participant gave some thought to reading level as a criterion. Dr. Lawler, for example, noted that one author was "not an easy person to read" (Interview 1). Some items were deemed too complex for students at the class level, while others were recognized as appropriate for graduate but not undergraduate students.

Perhaps the single most fascinating criterion was the demand that materials help engender a particular emotional and/or intellectual response in students. Rather than simply supplying factual data, the items chosen for class readings were also supposed to spark change and growth on an individual basis. Dr. Lawler, for example, needed to "help students think of American lesbians differently" (Interview 1). Likewise, Dr. Lewis described one item as "good for people who've never ever come in contact with African American women" (Interview 1). An interest in exposing people to new ideas, helping them understand the unfamiliar, and helping them recognize their own assumptions underlies a criterion that appeared repeatedly in work with all five participants.

Finally, five criteria centered on information characteristics: analytic or critical nature, depth of coverage, information format, perspective or viewpoint, and popular or scholarly nature. A few of these appear in the MacMullin and Taylor taxonomy of information traits (1984).

The analytic or critical nature of information was occasionally valued. Dr. Goodman, for example, liked items that "frame the issues" (Interview 1) and Dr. Lawler wanted an item that offered "contextualization" of an issue (Interview 2). This type of information was seen as central to the development of student understanding.

Varying depths of coverage were valued in different components of a course or in different types of courses. Dr. Lewis wanted one item on the grounds that it was "good for an undergraduate class since it's a survey" (Interview 2). On the other hand, Dr. Lawler considered one article "nice for the students to know about" but "too narrowly focused" for general reading; she determined that it would be added to the general bibliography for the course (Interview 2).

The information format criterion covered genres (e.g., poems, letters, diaries, statistics), format (e.g., list of subject headings), and physical structure (e.g., videotape, microfilm, newspaper). Dr. Taylor, for example, found bibliographies generally useful but much more so when annotated, especially when "working on someone you don't know who has a huge amount" written about her/him (Interview 1). Special issues of journals are "often

very good." Dr. Goodman would "assume there are a number of things in there" and "would definitely have put it on reserve" (Interview 1).

Perspective or viewpoint is an information characteristic that appeared frequently as a criterion. Although occasionally represented in some subject and indexing schemes, the substance of these judgements was seldom available to these participants in traditional document representations. Dr. Lewis wanted various views because "experiences have been obscured and distorted" (Interview 3). Dr. Lawler believed that "the definition of lesbian has too often ignored class—assumed a middle class, a white middle class, core to itself." She liked an article that specifically looked at class within lesbian movements (Interview 3).

The final information-characteristic criterion, popular or scholarly nature, is also indirectly included in some subject representations of documents. With terms such as "diaries" and "speeches" available, some citations indicate the primary or secondary nature of the material. Dr. Lawler valued both information types, noting, "Ideally every course would have primary sources as well as secondary" (Interview 2). Of course, the interest in primary materials sometimes overlaps with other criteria, particularly engendering a response in students. Dr. Jacobs, for example, likes congressional hearings, with their verbatim transcripts of women's testimonies, because they show the "actual voice of the activist" and are "as close as we can get in this environment to having these women speak in the classroom" (Interview 1).

Implications

Obviously a great deal more needs to be done on both this methodological approach and this research question. The approach is too labor intensive and time consuming for use in large-scale studies but it might well, after further case studies, be used to develop an instrument with which to identify and study the use of varying criteria involved in relevance decisions. Faculty reviewing printed citations may use criteria different from those used by others viewing citations online. The impact of the viewing mechanism merits examination. While these faculty requested that no Web sites be included in the citations located on their behalf, the rapid rise in the use of Web sites for instructional purposes would probably alter that request in future studies. How this methodology might be used to capture the relevance decisions made regarding Web sites is another issue.

Although these preliminary findings are extremely limited, they do indicate some useful contributions to the ongoing discussion of relevance. Information gathered, at least in part, to pass along to others may have unexpected criteria involved, such as emotional impact and readability. These criteria were not applied on a binary basis; faculty utilized three to five levels of utility ratings. Each classed certain items as immediately essential, immediately useless, and possibly useful. In at least one interview, each

participant also made other judgments such as "look at later, time permitting." Certainly the curiosity piqued by citations requires further study in the holistic context of personal information-seeking. Not only are needs situational and dynamic, but they are also active simultaneously and on different levels. Understanding the links between information needs, as revealed by the application of relevance criteria, would illuminate more than relevance alone.

References

Barry, C. L. (1994). User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science, 45*(3): 149–159.

Bookstein, A. (1979). Relevance. *Journal of the American Society for Information Science, 30*(5): 269–273.

Cool, C.; Belkin, N. J.; & Kantor, P. B. (1993). Characteristics of texts affecting relevance judgments. In *Proceedings of the 14th National Online Meeting*, M. E. Williams (Ed.). Medford, NJ: Learned Information, pp. 77–84.

Glaser, B. (1965). The constant comparative method of qualitative analysis. *Social Problems, 12*(4): 436–445.

Janes, J. (1991). Relevance judgements and the incremental presentation of document representations. *Information Processing & Management, 27*(6): 629–646.

Mizzaro, S. (1997). Relevance: The whole history. *Journal of the American Society for Information Science, 48*(9): 810–832.

Park, T. K. (1993). The nature of relevance in information retrieval: An empirical study. *Library Quarterly, 63*(3): 318–351.

Park, T. K. (1994). Toward a theory of user-based relevance: A call for a new paradigm of inquiry. *Journal of the American Society for Information Science, 45*(3): 135–141.

Regazzi, J. J. (1988). Performance measures for information retrieval systems: An experimental approach. *Journal of the American Society for Information Science, 39*(4): 235–251.

Saracevic, T. (1969). Comparative effects of titles, abstracts and full texts on relevance judgments. *Proceedings.* (Vol. 6, pp. 293–299.) Washington, D.C.: American Society for Information Science.